Published in final edited form as:

Phys Biol. ; 12(4): 045004. doi:10.1088/1478-3975/12/4/045004.

From Analog to Digital Models of Gene Regulation

Brian Munsky[1,2,#] and **Gregor Neuert**[3,4,#]

[1]Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO 80523, USA

[2]School of Biomedical Engineering, Colorado State University, Fort Collins, CO 80526, USA

[3]Department of Molecular Physiology and Biophysics, School of Medicine, Vanderbilt University, Nashville, TN, 37240, USA

[4]Department of Biomedical Engineering, School of Engineering, Vanderbilt University, Nashville, TN, 37240, USA

Abstract

Recently, major progress has been made to develop computational models to predict and explain the mechanisms and behaviors of gene regulation. Here, we review progress on how these mechanisms and behaviors have been interpreted with analog models, where cell properties continuously modulate transcription, and digital models, where gene modulation involves discrete activation and inactivation events. We introduce recent experimental approaches, which measure these gene regulatory behaviors at single-cell and single-molecule resolution, and we discuss the integration of these approaches with computational models to reveal biophysical insight. By analyzing simple toy models in the context of existing experimental capabilities, we discuss the interplay between different experiments and different models to measure and interpret gene regulatory behaviors. Finally, we review recent successes in the development of predictive computational models for the control of gene regulation behaviors.

Keywords

gene regulation; single-cell dynamics; biochemical noise; chemical master equation

1. Introduction

Life's diversity requires that every gene, within the simplest virus or the most complex nerve tissue, must control its expression to meet fluctuating biological demands. This task requires modulating the timing and strength of mRNA and protein production in response to different environments, various developmental stages, or other cellular processes. At a more basic level, each gene in each cell must decide how many messenger RNAs should be transcribed, and when. Obviously, to fully understand or predict the regulation of

[#]These authors contributed equally to this work and to whom correspondence should be addressed: munsky@engr.colostate.edu (B.M), gregor.neuert@vanderbilt.edu (G.N.)

PACS numbers: 02.50.Ey, 02.50.Ga, 02.50.Tt, 05.40.Ca, 82.39.-k, 82.39.Rt, 87.10.Mn, 87.15.Aa, 87.16.Ac, 87.16.dj, 87.18.Tt, 87.80.Nj

transcriptional kinetics is an enormous challenge, but one where single-cell experiments and stochastic analyses can help tremendously.

At the broadest level, gene regulation networks are described as vast collections of nodes connected by activating and repressive relationships [1]. When inferred from microarray [2, 3, 4], mass spectroscopy [5, 6, 7, 8], or DNA/RNA sequencing data [9, 10], these networks provide a high-level logical understanding of transcriptome regulation [11, 12]. However, more precise measurements suggest that regulatory inputs control gene expression at finer resolutions than simply ‘on’ or ‘off’. For a specific gene/regulator combination, one can measure a gene regulation function (GRF) to describe how regulatory input levels modulate transcriptional output levels [13]. Although the GRF can be measured at the population level using microarrays, Northern blotting, RT-PCR and RNA-Seq [14], single-cell measurements can reveal much richer regulatory behaviors [15, 16, 17, 18, 19]. In the following sections, we introduce some of these experimental tools, and we discuss a couple simple computational models that have been used to interpret their results. We also introduce a couple examples where different types of experiments alternately hide or reveal interesting temporal, spatial or stochastic gene regulatory response fluctuations, and we suggest model-based understanding could lead to the design of more informative experiments.

2. Single-cell Experiments

Many approaches [20, 21, 22] have been devised to uncover and quantify the temporal, spatial and cell-to-cell fluctuations in gene regulation at the single-cell level. The majority of these techniques rely upon natural or synthetic fluorescent reporters, such as genetically modified fluorescent proteins [23] or bioluminescent enzymes [24] as well as fluorescently labeled antibodies [25] and synthesized nucleic acids that hybridize to RNA [26, 27], all of which are measured with fluorescence microscopy, time-lapse microscopy [20, 24, 28, 29, 30], and flow cytometry [25, 31]. Many of these approaches provide a quantitative estimate of single-cell concentrations, but advanced single-molecule measurements can also directly quantify the precise number and intracellular locations of RNA [17, 26, 27, 32], protein [23, 33, 34] or both [28, 35]. In particular, the activation and deactivation of transcription is observed most directly through measurements of RNA, which can be accomplished with single-molecule RNA fluorescence in situ hybridization (smRNA-FISH), [17, 26, 27, 36, 37, 38], which allows one to measure directly the number and positions of RNA in fixed cells. Even when gene expression varies considerably from cell to cell, measurements of hundreds or thousands of individual cells enables precisely repeatable quantification of informative gene expression distributions [17, 39].

Although smRNA-FISH enables precise spatial measurements of individual endogenous RNA molecules, the technique requires fixation and permeablization and cannot provide more than a single temporal measurement per cell. At constant induction levels, this precludes the possibility of dynamic data needed to understand temporal fluctuations. To circumvent this constraint, one can take a number of approaches. By examining different populations of cells at different instances in time, endogenous *in situ* hybridization approaches can capture the spatiotemporal population dynamics of developmental processes [40, 41] or transient population responses to perturbations [17, 36, 42]. Conversely, *in vivo*

temporal measurements can be achieved by genetic modification of the transcribed mRNA to include a large number of MS2 bacteriophage hairpin structures and adding constitutively expressed MS2 coat proteins with GFP tags [29, 43, 44]. Under these modifications, when the modified target mRNA is expressed, it binds with GFP tags, and its movement throughout the cell can be captured with fluorescence time-lapse microscopy. Although this technique may perturb the endogenous mRNA dynamics, it allows direct estimates of many of the effects of the digital gene regulation model, including exponential waiting times between subsequent formations of active transcription sites [29] or between transcriptional bursts [28], exponentially distributed existence times for active transcription sites [28, 29], and geometrically distributed RNA burst sizes [28]. The MS2 approach technique has been extended to use similar interactions with PP7 bacteriophage hairpin structures, enabling two color real time analyses of two mRNA, simultaneously [45, 46].

In some cases, it is also possible to observe characteristics of the digital gene expression model at the protein level. By replacing or fusing endogenous mRNA with the coding region of a fluorescent protein, one can engineer a fluorescent output for a given gene of interest [47]. For DNA- or membrane-bound proteins, one can image individual protein molecules, and for freely diffusing protein, one can deconvolve the background fluorescence and calibrate to estimate the number of proteins per cell volume [34, 35]. Genetic modifications to introduce a luciferase-based assay [24, 48] can also introduce measurable fluorescence reporters of gene expression at the protein level. Analyses that monitor one or more spectrally distinct fluorescent signals with time-lapse experiments can lead to better understanding of the origins of variability and the causal connections between different regulatory proteins [49, 50, 51]. By combining fluorescent protein fusions with smRNA-FISH [35] or MS2-tag [28, 30] approaches, one can measure the single-cell mRNA and protein correlation. An interesting observation from such studies is that while the average mRNA and protein expression levels are correlated among different genes and different conditions, the single-cell numbers of mRNA and protein appear to be uncorrelated [35].

In the remainder of this article, we examine some of the computational tools and models that have been used to interpret, and in some cases predict, these experimental observations of gene regulation fluctuations from cell to cell and over time.

3. Analog and Digital Models of Gene Regulation

To introduce common gene regulatory behaviors and their interpretation, figure 1 illustrates two simple models of gene regulation. The ‘analog’ model allows direct, continuous tuning of the RNA transcription rate, whereas the ‘digital’ model has ‘off’ and ‘on’ states with fixed transcription rates. Where the analog model corresponds to a single gene state, the digital model [36, 37, 38, 52, 53] contains mutually exclusive ‘off’ and ‘on’ states between which genes transition with tunable rates k_{12} and k_{21} (figure 1(*a*)). Typically, the ‘off’ state has no transcription, while the ‘on’ state has a constant transcription rate, k_r. For either model, the translation rate is assumed to be proportional to the number of mRNA, and degradation obeys first-order reaction kinetics. While both the analog and digital models can produce the same average effects of gene regulation (figure 1(*b*)), the digital model is capable of capturing a much broader range of qualitatively different regulatory behaviors

over time (figure 1(*c*)) or from cell to cell (figure 1(*d–e*)). When detailed single-cell experiments capture these qualities, the analog model breaks down, but the digital model can help us to link regulatory behaviors to different physical mechanisms, such as chromatin conformation changes [37, 54, 55], DNA loop formations [23], transcription factor binding configurations [24, 36, 56], formation of nuclear transcription factories [57] or other discrete effects.

The digital gene expression model has been analyzed with numerous computational methods, including kinetic Monte Carlo simulations [58], moment generating function approaches [52, 53, 59], random telegraph analyses [24, 37, 38], Laplace transforms [60], and finite state projection (FSP, [17, 61]) approaches. For the analyses in this work, we utilize a modification of the FSP analysis [61]. This approach enables us to solve directly for the joint probability distributions of gene state, mRNA level and protein level at any time t_2 given an initial distribution at a previous time t_1. Using this analysis, figure 2 illustrates some of the most important regulatory behaviors of the models for a few different phenotypical cases. Figure 2(*a*) plots representative time-trajectories of a single-cell's gene-state and its mRNA and protein levels. Here, 'on' and 'off' intervals are exponentially distributed with means $1/k_{21}$ and $1/k_{12}$, respectively, and the number of mRNA produced per 'on' interval (or burst size) is geometrically distributed with mean k_r/k_{21} [28]. Statistics of cell-to-cell fluctuations are captured in figure 2(*b–c*), which plots the equilibrium marginal and joint probability distributions for mRNA and protein molecules. Each of the regulatory behaviors shown in figure 2 depends heavily on the relative time scales of gene transitions and the degradation rates of mRNA and protein (different columns of figure 2). When gene transitions are fast compared to mRNA degradation (left column), the mRNA trajectories and distributions take on the characteristics of a constitutively expressed gene and are indistinguishable from the analog model [16]. However, when gene transition rates are comparable with mRNA decay rates (middle and right column), mRNA pulsing becomes apparent (figure 2(*a*), middle and right), and mRNA distributions take on bimodal characteristics (figure 2(*b*)). However, proteins only inherit this bimodality when protein degradation is fast [59] (compare middle and right columns).

In many cases, RNA measurements provide quantitative insight that supports the digital model of gene regulation. For example, the analog model predicts mRNA to exhibit a Poisson distribution (figure 1(*e*), left), in which the variance σ^2 is equal to the mean expression μ and the Fano factor, $F = \sigma^2/\mu$, is unity [62]. However, many genes exhibit higher Fano factors [28, 35, 37, 38], as predicted by the digital transcription model. In some cases, spatial information from smRNA-FISH experiments also supports the digital gene expression model by revealing that some cells have large groups of transcripts at transcription sites [32, 36] or in the nucleus, and other cells contain mostly cytoplasmic mRNA [37]. Such observations are also consistent with the digital model's prediction of intermittent active and inactive transcriptional periods [37]. However, in contrast to direct RNA measurements, it is a little more difficult to discriminate between digital and analog gene regulatory models by monitoring long-lived protein reporters. Since protein levels inherit and increase the variability of mRNA [63], their variance is typically non-Poisson and may be bimodal even for the analog gene expression model.

To explore differences in the statistics of mRNA and protein fluctuations and to explain the missing mRNA-protein correlation observed in [35], it is helpful to examine numerically the expected temporal single-cell fluctuations of mRNA and protein. Using FSP analyses [61] for the cases introduced in figure 2, we can compute the joint probability distributions of mRNA at time, t, and protein at time, $t + \tau$ (figures 3(*b*,*c*)). Further, we can quantify the cross-correlations between mRNA (m) and protein (p) for all possible values of τ. Using this analysis, figure 3(*c*) plots the cross-correlation function, $R_{mp}(\tau)$, defined as:

$$R_{mp}(\tau)=\frac{\mathbb{E}\{(m(t)-\mu_m)(p(t+\tau)-\mu_p)\}}{\sigma_m\sigma_p} \quad (1)$$

as a function of the time lag ($\tau > 0$) or lead ($\tau < 0$) of protein, $p(t + \tau)$, with respect to mRNA, $m(t)$. Depending upon the model parameters, the magnitude of $R_{mp}(\tau)$ can range from one (perfect correlation) to zero (no correlation). A close inspection of figure 3 provides two insights to understand the observed absence of mRNA-protein correlations at the single-cell level [35]. First, we note that many mRNA degrade faster (5–10 minutes) compared to fluorescent protein reporters (30 minutes or more) [35]. In figure 3(*b*,*c*) (middle column), we see that fast mRNA dynamics and slow protein dynamics strongly diminishes the level of mRNA-protein correlation. In this case, proteins exist long after their corresponding mRNA has degraded, and the instantaneous protein level depends upon a long time integral of the mRNA level, not the instantaneous mRNA level. Conversely, for mRNA with comparatively long lifetimes, instantaneous mRNA and protein levels are much better correlated [28] (figure 3, left and right). Second, figure 3(*c*) shows that the cross-correlation function is not symmetric about $\tau = 0$. Since protein translation lags behind mRNA transcription, the maximum of occurs at a positive value of τ, and the correlation drops off quickly as τ decreases. Although smRNA-FISH captures both nascent and mature mRNA [37, 36], fluorescent protein reporters capture only those proteins that have completed translation, folding, and chromophore oxidation. These processes take two minutes in vitro for the fastest available YFP variant [64] and typically 5–30 minutes or more for other fast-maturing fluorescent proteins [64, 65]. With such folding times, a snapshot of protein and mRNA levels effectively measures the cross-correlation between mRNA at time, t, and protein levels at the earlier time $t - \tau_m$, where maturation introduces the apparent delay τ_m. Figure 3 illustrates the effect that typical maturation times (τ_m= 5 and 30 minutes, respectively) have on measured correlations between mRNA and protein. Clearly, measurable correlations between mRNA and protein can be very sensitive to the fluorescent protein maturation time as well as the protein life times. To some extent, these time-scale concerns can be reduced through additional genetic modifications that introduce a luciferase-based assay [24, 48], which combines a much shorter protein lifetime with a faster fluorescence reporter. The time derivative of this luciferase intensity is correlated with the mRNA level, which allows one to infer quantitative values for the live cell mRNA bursting dynamics [24].

4. Discussion

With the computational and experimental tools described above, we can now discuss how different genes regulate their behaviors. In general, housekeeping genes exhibit fewer

bursting behaviors [32] and less correlation between mRNA of different types [66], both of which are consistent with analog models of regulation. Conversely, mRNA from stress response genes and temporally induced genes exhibit much greater bursting behaviors and are more likely to be correlated [66]. These observations argue that the main function of multiple gene states is to enable genes to modulate their expression in response to environmental and intracellular cues. In principle, such regulation could be achieved through any positive or negative modification of one or more rate parameter. For example, reactions could correspond to the direct binding of activators or repressors, respectively, in which case the rates k_{12} or k_{21} would increase with the transcription factor concentration. Indirectly, regulators could inhibit transitions by sequestering or modifying other bound or unbound transcription factors. Alternatively, regulators could stabilize or destabilize mRNA or proteins to modify their effective degradation rates. Additional effects could correspond to chromatin modifications [38, 54, 55], polymerase recruitment [41, 57], DNA loop formations [23], or some other gene regulatory process. Because certain genes may be controlled by the outputs of other genes or signaling processes, regulatory inputs themselves are often dynamic signals with their own frequency and amplitude modulations [67, 68].

Figure 4 illustrates how different gene regulatory mechanisms would respond to temporal changes in an inducer, $x(t)$, such as a varying transcription factor, chromatin modifier or signal transduction protein concentration. We consider four possible models (figure 4(*a*)), including the analog model ($\hat{k_r} = k_r(x+cx)/(c+x)$, black) and three digital models, where a single reaction is affected by the inducer according to: $\hat{k_{12}} = k_{12}x$ (gold), $\hat{k_{21}} = k_{21}/x$ (magenta), or $\hat{k_r} = k_r(x + cx)/(c + x)$ (cyan). Here $c = k_{21}/k_{12}$, and all parameters are chosen such that all four models predict the same equilibrium average for any finite $x > 0$. Each model is subjected to a moderate (10 fold) and large (1000 fold) step increases in inducer as shown in figure 4(*b*), left and right respectively. The corresponding dynamics of the average regulatory behaviors are shown in figure 4(*c*). Representative mRNA distributions before, during, and after the induction response are shown in figure 4(*d*), and effects on the Fano factor are shown in figure 4(*e*) and 4(*f*). Despite identical steady-state averages, the models produce different temporal dynamics and/or cell-to-cell fluctuations. Figure 4 shows that the three digital models predict different dynamics (compare gold, magenta and cyan lines), but the average dynamics of the analog model and the digital k_r modulation model are identical (compare gold and black lines). At the low induction level, $x = 1$, all digital models predict the same steady state distribution (see figure 4(*d*) at 0 hr), but these distributions are distinct at moderate induction levels (see figure 4(*d*), left at 10 hr). At very high induction levels, one can make some interesting comparisons between the analog model and the different versions of the digital model (see figure (*b–f*) right). First, when subjected to high induction, the analog model and the k_{12}-modulated models are indistinguishable in both average dynamics (see figure 4(*c*) right, where the black and magenta lines overlap) and induced single-cell distributions (see figure 4(*d*) right at 10 hr). Second, the analog and k_{21}-modulated models match the high-induction equilibrium distributions (see figure 4(*c*) right at 10 hr, where the black and cyan lines overlap), but not the transient dynamics (see figure 4(*b*) right, where the black and cyan lines are distinct). Finally, the analog and k_r-modulated model match only the average transient dynamics (see figure 4(*c*), where the black and gold lines overlap), but not the equilibrium single-cell distributions (see figure 4(*d*), where the

black and gold lines are distinct). In other words, at high induction one can distinguish the k_r-, k_{21} and k_{12}-modulated models from the analog model, but to do so requires different types of data (i.e., distributions for k_r, transients for k_{21} and both simultaneously for k_{12}). Since our intention is to probe understanding for the mechanisms of gene regulation, the fact that different hypotheses require different experiments suggests that experiments and models should be designed in concert with one another.

Because different models predict different dynamics and/or cell-to-cell fluctuations, it follows that well-chosen single-cell measurements at different conditions or times can help distinguish between competing mechanisms [17, 23, 37, 38, 54, 55, 69]. In the case of induction of PHO5 in budding yeast, the relative intrinsic noise at high expression levels is reduced [54, 55]. This trend was attributed to slow gene activation, and modulation of the gene transitions [54, 55], a trend that is similar to that plotted by the gold and cyan curves of figure 4(*f*). smRNA-FISH analysis of the *tet* promoter in mammalian cells shows that transcriptional activators have strong effects specifically on mRNA burst size through k_{21} or k_r, but not on the frequency controlled by k_{12} [37, 38]. Similar observations made for the *lac* operon in *E. coli* showed that lengths of bursts and not the frequencies were controlled by inducer concentrations [23]. This trend, in which the reaction k_{21} is controlled by the inducer, remained consistent for an additional 20 genes in *E. coli* as measured with smRNA-FISH and MS2-tag analyses [38]. However, the trend of k_{21}-modulated control is not absolute, and modulation of the parameter k_{12} has been shown to be consistent with other gene regulatory systems [36]. At present, it is not clear which biological circumstances would call for one mechanism over the other. One could speculate that since k_{12}-modulation allows for faster and less variable responses to extracellular cues, it may be an advantageous strategy for tightly regulated genes in intercellular signaling or developmental pathways within multi-cellular organisms. On the other hand, less-sensitive k_{21}-modulated control may be better suited to reject fast environmental fluctuations and enable population diversity for single-cell organisms in changing environments.

Furthermore, although the two-state model quantitatively captures bursting regulatory behavior for many gene regulatory processes, it remains an approximation of more complicated physical systems. In some cases, dynamic single-cell measurements can reveal more aspects of this complexity [17, 23, 24, 29, 44]. In the social amoeba *dictyostelium*, induction response of the *discoidin Ia* gene exhibits responding and non-responding phenotypes, in which responding cells are characteristic of a two-state system, and non-responding phenotypes exhibit slow transitions characteristic of an additional ‘off’ state [29]. Non-exponentially distributed times between activation events have also suggested an additional inactive state for genes in mammalian cells [24]. In *E. coli*, combinations of inactive periods, frequent small bursts, and rare large bursts suggest three different states corresponding to ‘off’, ‘leaky’ and ‘on’ for lac expression [23]. RNA-FISH observations for many genes in *drosophilla* embryos also suggest common models of ‘off’, ‘poised-polymerase’ and ‘on’ states, in which certain genes containing stalled Pol II preference the ‘poised’ state and respond faster and less variably than those without [41]. Furthermore, the loading of Pol II can vary between multiple different ON states [44]. In each of these

studies, insight beyond the two-state model requires dynamic data, either from live cell tracking or from multiple snapshots of fixed cells during a dynamic response [17].

The merger of experimental and computational analyses enables systematic testing and refinement of understanding for gene regulation. Here we have focused on how single-cell approaches and dynamic measurements can reveal many qualitative features in which this regulation behaves as one predicts using a digital model. In time, these techniques will also enable quantitative predictions for regulatory behaviors in new and untested conditions, such as new experimental conditions or for new regulatory constructs [17]. Such predictive capabilities will eventually enable more precise analysis and manipulation of biomedical processes as well as more reliable tools to design synthetic constructs to meet bioengineering challenges.

Acknowledgments

B.M. was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R25GM105608. G.N. was supported by the National Institutes of Health under award number DP2GM114849.

References

1. Breker M, Schuldiner M. Nature Reviews Molecular Cell Biology. 2014; 15:453–464. [PubMed: 24938631]
2. Chuang HY, Hofree M, Ideker T. Annual Review of Cell and Developmental Biology. 2010; 26:721–744.
3. Cahan P, Li H, Morris SA, Lummertz da Rocha E, Daley GQ, Collins JJ. Cell. 2014; 158:903–915. [PubMed: 25126793]
4. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR. Science (New York, NY). 2006; 313:1929–1935.
5. Cox J, Mann M. Annual Review of Biochemistry. 2011; 80:273–299.
6. Swaney DL, Beltrao P, Starita L, Guo A, Rush J, Fields S, Krogan NJ, Villén J. Nature Methods. 2013; 10:676–682. [PubMed: 23749301]
7. Bensimon A, Heck AJR, Aebersold R. Annual Review of Biochemistry. 2012; 81:379–405.
8. Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmair A, Faerber F, Kuster B. Nature. 2014; 509:582–587. [PubMed: 24870543]
9. Meyer CA, Liu XS. Nature Reviews Genetics. 2014; 15:709–721.
10. Kiel C, Beltrao P, Serrano L. Annual Review of Biochemistry. 2008; 77:415–441.
11. Carvunis AR, Ideker T. Cell. 2014; 157:534–538. [PubMed: 24766803]
12. Shen-Orr SS, Milo R, Mangan S, Alon U. Nature Genetics. 2002; 31:64–68. [PubMed: 11967538]
13. Hao N, Budnik BA, Gunawardena J, O'Shea EK. Science. 2013; 339:460–464. [PubMed: 23349292]
14. Albeck JG, MacBeath G, White FM, Sorger PK, Lauffenburger DA, Gaudet S. Nature Reviews Molecular Cell Biology. 2006; 7:803–812. [PubMed: 17057751]
15. Raj A, van Oudenaarden A. Cell. 2008; 135:216–226. [PubMed: 18957198]
16. Munsky B, Neuert G, van Oudenaarden A. Science. 2012; 336:183–187. [PubMed: 22499939]
17. Neuert G, Munsky B, Tan RZ, Teytelman L, Khammash M, van Oudenaarden A. Science. 2013; 339:584–587. [PubMed: 23372015]
18. Hoppe PS, Coutu DL, Schroeder T. Nature Cell Biology. 2014; 16:919–927. [PubMed: 25271480]

19. Stasevich TJ, Hayashi-Takanaka Y, Sato Y, Maehara K, Ohkawa Y, Sakata-Sogawa K, Tokunaga M, Nagase T, Nozaki N, McNally JG, Kimura H. Nature. 2014
20. Locke JCW, Elowitz MB. Nature Reviews Microbiology. 2009; 7:383–392. [PubMed: 19369953]
21. Bao G, Rhee WJ, Tsourkas A. Annual Review of Biomedical Engineering. 2009; 11:25–47.
22. Coulon A, Chow CC, Singer RH, Larson DR. Nature Reviews Genetics. 2013; 14:572–584.
23. Choi PJ, Cai L, Frieda K, Xie XS. Science. 2008; 322:442–446. [PubMed: 18927393]
24. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Science. 2011; 332:472–474. [PubMed: 21415320]
25. Perfetto SP, Chattopadhyay PK, Roederer M. Nature Reviews Immunology. 2004; 4:648–655.
26. Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. Nature Methods. 2008; 5:877–879. [PubMed: 18806792]
27. Femino AM, Fay FS, Fogarty K, Singer RH. Science. 1998; 280:585–590. [PubMed: 9554849]
28. Golding I, Paulsson J, Zawilski SM, Cox EC. Cell. 2005; 123:1025–1036. [PubMed: 16360033]
29. Chubb JR, Trcek T, Shenoy SM, Singer RH. Current Biology. 2006; 16:1018–1025. [PubMed: 16713960]
30. Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. Science. 2011; 332:475–478. [PubMed: 21512033]
31. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, Barkai N. Nature Genetics. 2006; 38:636–643. [PubMed: 16715097]
32. Zenklusen D, Larson DR, Singer RH. Nature Structural & Molecular Biology. 2008; 15:1263–1271.
33. Cai L, Friedman N, Xie XS. Nature Cell Biology. 2006; 440:358–362.
34. Yu J, Xiao J, Ren X, Lao K, Xie XS. Science (New York, NY). 2006; 311:1600–1603.
35. Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS. Science. 2010; 329:533–538. [PubMed: 20671182]
36. Senecal A, Munsky B, Proux F, Ly N, Braye FE, Zimmer C, Mueller F, Darzacq X. Cell Reports. 2014; 8:75–83. [PubMed: 24981864]
37. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. PLoS Biology. 2006; 4:1707–1719.
38. So, Lh; Ghosh, A.; Zong, C.; Sepúlveda, LA.; Segev, R.; Golding, I. Nature Genetics. 2011; 43:554–560. [PubMed: 21532574]
39. Schwabe A, Bruggeman FJ. Nature Communications. 2014; 5:4798.
40. Raj A, Rifkin SA, Andersen E, van Oudenaarden A. Nature. 2010; 463:913–918. [PubMed: 20164922]
41. Boettiger AN, Levine M. Science. 2009; 325:471–473. [PubMed: 19628867]
42. Tan RZ, van Oudenaarden A. Molecular Systems Biology. 2010; 6:1–7.
43. Park HY, Lim H, Yoon YJ, Follenzi A, Nwokafor C, Lopez-Jones M, Meng X, Singer RH. Science. 2014; 343:422–424. [PubMed: 24458643]
44. Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, Levine M. Proceedings of the National Academy of Sciences. 2014; 111:10598–10603.
45. Hocine S, Raymond P, Zenklusen D, Chao JA, Singer RH. Nature Methods. 2013; 10:119–121. [PubMed: 23263691]
46. Coulon A, Ferguson ML, de Turris V, Palangat M, Chow CC, Larson DR. eLife. 2014:3.
47. Shaner NC, Steinbach PA, Tsien RY. Nature Methods. 2005; 2:905–909. [PubMed: 16299475]
48. Norris AJ, Stirland JA, McFerran DW, Seymour ZC, Spiller DG, Loudon ASI, White MRH, Davis JRE. Molecular Endocrinology. 2003; 17:193–202. [PubMed: 12554747]
49. Dunlop MJ, Cox RS, Levine JH, Murray RM, Elowitz MB. Nature Genetics. 2008; 40:1493–1498. [PubMed: 19029898]
50. Pedraza JM, van Oudenaarden A. Science. 2005; 307:1965–1969. [PubMed: 15790857]
51. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. Molecular Systems Biology. 2007; 3:1–4.
52. Ko MS. Journal of Theoretical Biology. 1991; 153:181–194. [PubMed: 1787735]

53. Peccoud J, Ycart B. Theoretical Population Biology. 1995; 48:222–234.
54. Raser JM, O'Shea EK. Science. 2004; 304:1811–1814. [PubMed: 15166317]
55. Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EEM. Nature. 2014; 512:96–100. [PubMed: 25043061]
56. Bumgarner SL, Neuert G, Voight BF, Symbor-Nagrabska A, Grisafi P, van Oudenaarden A, Fink GR. Molecular Cell. 2012; 45:470–482. [PubMed: 22264825]
57. Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, Dugast-Darzacq C, Hajj B, Dahan M, Darzacq X. Science. 2013; 341:664–667. [PubMed: 23828889]
58. Gillespie DT. Journal of Physical Chemistry. 1977; 81:2340–2361.
59. Iyer-Biswas S, Hayot F, Jayaprakash C. Physical Review E. 2009; 79:031911.
60. Friedman N, Cai L, Xie X. Physical Review Letters. 2006; 97:168302. [PubMed: 17155441]
61. Munsky B, Khammash M. Journal of Chemical Physics. 2006; 124:044104. [PubMed: 16460146]
62. Thattai M, van Oudenaarden A. Proceedings of the National Academy of Sciences. 2001; 98:8614–8619.
63. Paulsson J. Nature. 2004; 427:415–418. [PubMed: 14749823]
64. Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. Nature Biotechnology. 2002; 20:87–90.
65. Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, Tsien RY. PNAS. 2002; 99:7877–7882. [PubMed: 12060735]
66. Gandhi SJ, Zenklusen D, Lionnet T, Singer RH. Nature Structural & Molecular Biology. 2010; 18:27–34.
67. Cai L, Dalal CK, Elowitz MB. Nature. 2008; 455:485–490. [PubMed: 18818649]
68. Tay S, Hughey JJ, Lee TK, Lipniacki T, Quake SR, Covert MW. Nature. 2010; 466:267–271. [PubMed: 20581820]
69. Munsky B, Trinh B, Khammash M. Molecular Systems Biology. 2009; 5:1–7.

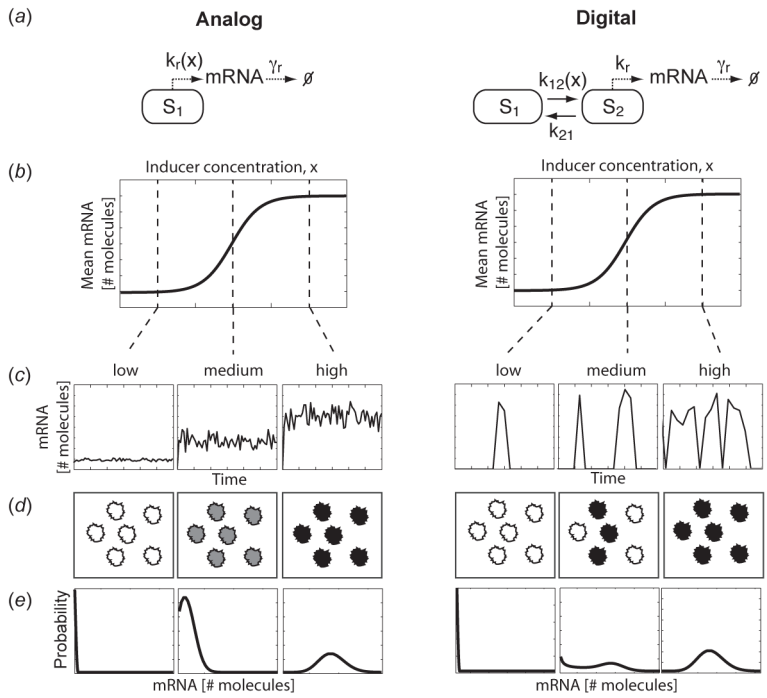

Figure 1. Analog and Digital Models of Gene Regulation

(*a*) Schematics of analog and digital regulation models. (*b*) Both models may yield equivalent average (bulk) mRNA changes versus inducer level. (*c*) Single-cell temporal trajectories for mRNA populations at low medium and high induction levels. (*d*) Single-cell behaviors at different expression levels: in the analog model, expression level is modulated in every cell; in the digital model, the number of responding cells may also be modulated. (*e*) Probability distributions for the mRNA populations at different inducer concentrations.

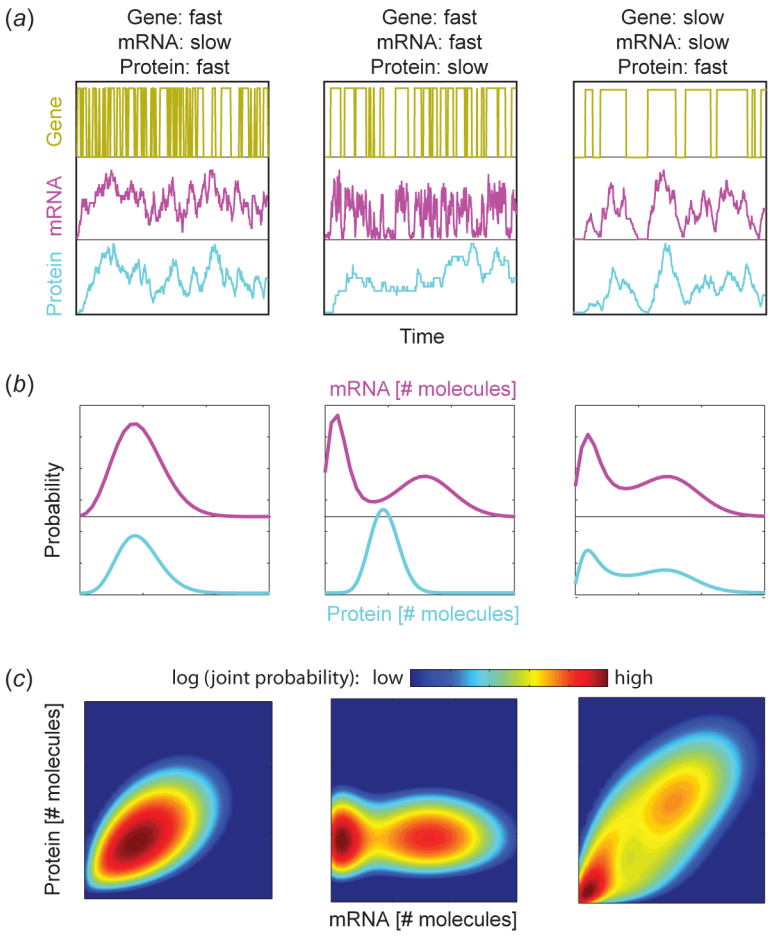

Figure 2. Fluctuations of the Digital Gene Expression Model

To illustrate different regulatory behaviors of the digital gene expression model, three parameter regimes with different relative speeds of gene state, mRNA and protein fluctuations are examined. Left: $\{k_{12} = k_{21} = 0.33\ \text{min}^{-1},\ \gamma_r = 0.067\ \text{min}^{-1},\ k_r = 1\ \text{min}^{-1},\ \gamma_p = 0.1\ \text{min}^{-1},\ k_p = 0.5\ \text{min}^{-1}\}$. Middle: $\{k_{12} = k_{21} = 0.0167\ \text{min}^{-1},\ \gamma_r = 0.067\ \text{min}^{-1},\ k_r = 1\ \text{min}^{-1},\ \gamma_p = 0.1\ \text{min}^{-1},\ k_p = 0.5\ \text{min}^{-1}\}$ Right: $\{k_{12} = k_{21} = 0.1\ \text{min}^{-1},\ \gamma_r = 1\ \text{min}^{-1},\ k_r = 15\ \text{min}^{-1},\ \gamma_p = 0.011\ \text{min}^{-1},\ k_p = 0.056\ \text{min}^{-1}\}$ These parameter changes lead to distinctly different regulatory behaviors at several levels: (*a*) Temporal trajectories for gene state (green), mRNA numbers (red), and protein numbers (blue). (*b*) Marginal mRNA and protein distributions at equilibrium. (*c*) Joint mRNA and protein distributions at equilibrium.

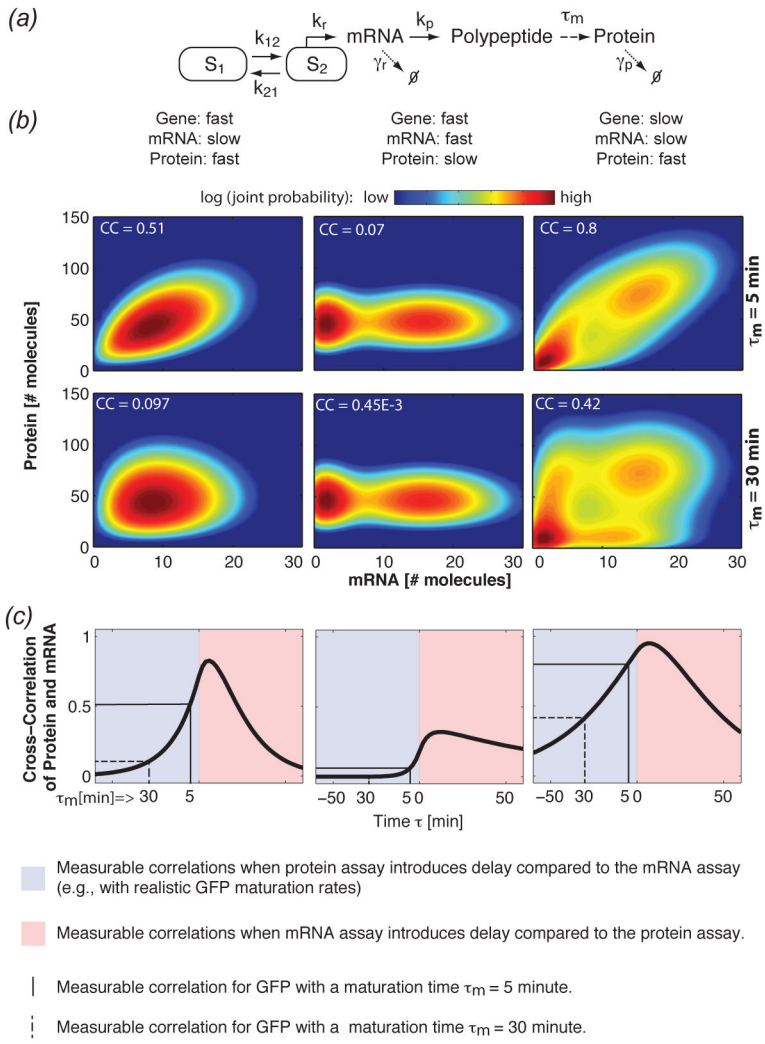

Figure 3. Cross-Correlations of the Digital Model

Time scales of gene, mRNA and protein fluctuations have strong impacts on temporal correlations and sensitivity to measurement delays. Parameters are as given in figure 2. (*a*) Simple model of gene transcription, mRNA translation and protein maturation with fixed time delay (τ_m). (*b*) Model predicted joint probability distribution of mRNA and protein assuming a fast protein maturation time of τ_m = 5 minutes (top row) and a moderate protein maturation time of τ_m = 30 minutes (bottom row). CC corresponds to the equilibrium cross-correlation of mRNA at time *t* and protein at time $t - \tau_m$. (*c*) Equilibrium cross-correlation between mRNA and protein, $R_{mp}(\tau)$, versus time delay, τ. Vertical lines represent maturation times τ_m of 5 (solid) and 30 (dashed) minutes.

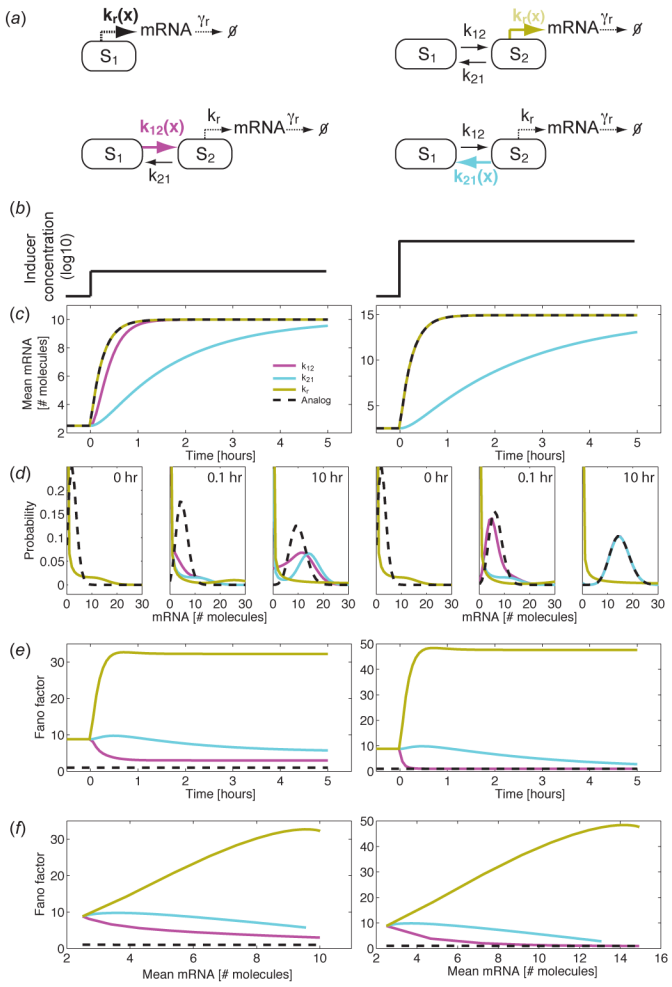

Figure 4. Effects of Induction Changes

(*a*) Induction levels, $x(t)$, affect the analog model according to $\hat{k_r} = k_r(x + cx)/(c + x)$ (black), and the digital model according to $\hat{k_{12}} = k_{12}x$ (magenta), $\hat{k_{21}} = k_{21}/x$ (cyan), or $\hat{k_r} = k_r(x + cx)/(c + x)$ (gold). (*b*) Inducer level, $x(t)$, undergoes a step increase at $t = 0$ from $x = 1$ to either $x = 10$ (left) or $x = 1000$ (right). (*c*) Average mRNA dynamics for each model. (*d*) Distributions before, during, and after induction change. (*e*,*f*) Fano factor as a function of time (*e*) and mean expression (*f*). For the analog model, $F(t) = 1$, and for the digital models, $F(t) > 1$. Parameters used for this illustration are: $\{k_r = 1\ \text{min}^{-1}, k_{12} = 0.0067\ \text{min}^{-1}, k_{21} = 0.033\ \text{min}^{-1}, \gamma = 0.067\ \text{min}^{-1}, c = k_{21}/k_{12}\}$.s